

HMMER Sequences in the Env Alignment

A_U455	HIVU455	M62320	Oram,J.D.	ARHR 6, 1073 (1990)
B_HXB2R	HIVHXB2R	K03455	Starcich,B.	Science 227, 538 (1985)
C_UG268A2	HIVUG268A2	L22948	Louwagie,J.J.	JVI 69, 263 (1995)
D_ELI	HIVELI	K03454	Alizon,M.	Cell 46, 63 (1986)
E_TN2432	HIVTN2432	L03703	McCutchan,F.E.	ARHR 8, 1887 (1992)
F_BZ163A	HIV1BZ163A	L22085	Louwagie,J.J.	ARHR 10, 561 (1994)
G_LBV217	HIVLBV217	L11778	Louwagie,J.J.	AIDS 7, 769 (1993)
O_ANT70C	HIVANT70C	L20587	Vanden Haesevelde,M.	JVI 68, 1586 (1994)
O_MVP5180	HIVMVP5180	L20571	Gurtler,L.G.	JVI 68, 1581 (1994)
O_VAU	HIV1VAU	X80020	Charneau,P.	Virology 205, 247 (1994)
CPZGAB	SIVCPZGAB	X52154	Huet,T.	Nature 345, 356 (1990)
A_ROD	HIV2ROD	M15390	Clavel,F.	Nature 324, 691 (1986)
B_EHOA	HIV2EHOA	U27200	Rey-Cuille,M.A.	Virology 202, 471 (1994)
SD_MM251	SIVMM251	M19499	Franchini,G.	Nature 328, 539 (1994)
STM_STM	SIVSTM	M83293	Novembre,F.J.	Virology 186, 783 (1992)
VER_AGM3	SIVAGM3	M30931	Baier,M.	Virology 176, 216 (1990)
GRI_AGM677	SIVAGM677	M66437	Fomsgaard,A.	Virology 182, 397 (1991)
SYK_SYK	SIVSYK	L06042	Hirsch,V.M.	JVI 67, 1517 (1993)

The following alignment and most-likely sequence were generated by the HMMER program as described in Part III. Approximately 400 HIV and SIV sequences contributed to the output; for simplicity, only representative subtypes and types are shown. Hybrid (i.e. recombinant) viral sequences will not affect the profile. The annotation was based upon information for HIV-1s and therefore the user should be cautious about its applicability to other PIV sequences. Cysteines are indicated by '**'. Potential N-linked glycosylation sites are indicated by '^^^'. Four highly conserved residues involved in CD4 interaction have been annotated. Eight conserved, gapless blocks (shaded) were independently derived from the representative sequences below (excluding the most-likely sequence) using the BLOCKMAKER and Motif programs; in most instances, the boundaries of these blocks are consistent with the HMM alignment. Block sizes are limited to 55 residues, hence contiguous blocks may occur. Five additional blocks over the C-terminal region of gp41 were separately determined, as variability and the existence of prematurely truncated (or unsequenced) representatives necessitated a separate determination. These five blocks should be given less weight. The Motif "Cobbler" sequence (involving just the first eight blocks) follows this alignment.

	-> vpU signal peptide / gp120 *	
most-likely	MR.V....KGIRRNYQHLWRW.....GI....LLLGMMLI.CS.AAENLWVTVYYGVPVWKE	46
A_U455	--.----M--Q---PC-----T...MI--LII..-N.-.QQ-----D	45
B_HXB2R	--.----E...K-----WRWGTM-----T-K-----	47
C_UG268A2	--.----M--Q--C-QW-I-----GFWI---.N.VMG-----	46
D_ELI	--.A....R--E--C-NW-K-----M---I--T---D-----	46
E_TN243	--.----ETQM-WPN--K-----T....I--LVI----SD-----RD	46
F_BZ163A	--.----R-MQ--W--GK-----L....F--I-L--N.-.-----D	45
G_LBV217	--.A.....L--W---K-----I--LVI--N.-SG-----A-ED	46
O_ANT70	-I.-TMKAME.K--.KK-----TL....Y-AMA-ITP-L.SLRQ-YA--A---ED	47
O_MVP5180	-T.-T..M-VMKK-NRKS-----SL....YIAMA-L-P-LSYSKQ-YA--S---E-	48
O_VAU	-T.AIMKAM--K--.RK-GI-----CL....I-ALIIPC.L-.C.NQ-YA--S---ED	47
CPZGAB	-K.-....MEKKKR.....DWNS.....LSIIT....IITII-LTP-L.TS--.-----HD	46
CPZANT	--.K.....PIHII-----L....A--IQFIE.KG.TN--DY---F-----RN	38
A_ROD	-M.N.....-L-I-----A-----ASACLV.Y..C.T.QY--F---T--N	34
B_EHOA	-AH-----NYL-V-----T-LLIS-YG.YMGKNF--F--I-A--N	37
SD_MM251	-G.C.....LG--.L-I-----A-----SVYG-Y..C.T.QY--F---A-RN	37
STM_STM	-A.C.....PG--.L-I-----A-----SACLT.Y..C.T.QY--F---A-RN	37
VER_AGM3	-K.L.....TL-I-----I--IGVV.LN.T.RQQ--F-----N	36
GRI_AGM677	-G.R.....L-I-----K....I-IIAIG-SI.GIG--Y--F--I---N	35
TAN_AGMB14	XQ.A....FCMTP-----	9
SAB_SAB1C	-K.L.....LT....V--WLSCG.W-LVWLVQY--F--I---N	34
SYK_SYK	-A.A.....FRTYIVCLFLSISLGF.....MEKQQY--F--I-H-ED	37
* * ^^^		
most-likely	ATTTLFCASDAKAYDTENVNVWATHACVP..TDP..NPQ.EIVLENVT.ENFNMW..K..NN.M....V.EQM	103
A_U455	-V-----A-----D-V--.E-----D--	102
B_HXB2R	-----V-----V--V-----D-----	104
C_UG268A2	-----E-----S-----D-----D--	103
D_ELI	-----S-E--A--I-----A-----	103
E_TN243	-D-----HE-----Y-----N-----	103
F_BZ163A	-----S-ER-----V-----D-----S-----	102
G_LBV217	-D-----S-S-----S-N-----	103
O_ANT70	--PV-----NLTS--K--I--SQ-----T-Y--YP-H--DD--I-----Y-----	104
O_MVP5180	-APV-----NLTS--Q--I--SQ-----H--FP-G--D-DI-----Y-----D--	105
O_VAU	-KP-----NLTS--Q--I--Q-----S-N.-YE-K--.GK--I--.Y.I-----D--	104
CPZGAB	-DPV-----HS--A--I--Q-----S--.VF-P--I.-S-----D--	103
CPZANT	--P-----TN-SMTS-----TS--.I--.D-I.VVR--.TS.VW--AY-----S-	94
A_ROD	--IP----TR.NR.....D.T-G-IQ--L..P-N..DDYQ--T--.A-DA--.N....T....-T--A	83
B_EHOA	-SIP----TR.NR.....D.T-G-VQ--L..P-N..DDYT--Q--.I--.A-DA--.D....T....-TD-A	86
SD_MM251	--IP----TK.NR.....D.T-G-TQ--L..P-N..GDYS-LA--.S-DA--.E....T....-T--A	86
STM_STM	--IP----TK.NR.....D.T-G-TQ--L..P-N..GDYS-LAI--.A-DA--.D....T....-T--A	86
VER_AGM3	SSVQA--MTPTR.....L---TNS.I..P-D..HDYT-VP--I--P-EA-ADR..P.L....A-A	87
GRI_AGM677	S-VQA--MTP.NT.....M---TN-I--D-H..DNT..VP--I--A-EA--D..P.L....K-A	84
TAN_AGMB14T-L---TN-I--D-H..-Y-.VQ--.S--K-EA--.DR-P.L....A-A	49
SAB_SAB1C	SSVQA--KTP.NT.....L--STN-I--D-EPEGTIA-VPIP-I--K-DA--.RNPL-.G-A	89
SYK_SYK	-YAP--TTSH-G.....G--KN--SA.....D-I-VRV.-I-G-Y-PA--N..SSH--.IRQ-I	88

HIV1 ENV

	<-----	V1 Loop	
	* * * ^ ^	^ ^ ^ ^ ^	^ ^ ^ *
most-likely	HEDIISLWDQSLKPCVKLTPLCVTLNCTDV.....	NAT.N.TNNTTNNTKIDMI..NETSSCIRQ	159
A_U455	-----D-HNIT.....	I-N--NNT-I-DGV.....	147
B_HXB2R	-----S-K--LK.....	D--.---SSSGRM.....	148
C_UG268A2	-Q-V-----N-N.....	V-I--NA-A--SPYE.....	149
D_ELI	-----S-ELR.....	N-G..M.G--V-T.....	146
E_TN243	Q--V-----I-NAKLT.....	NA-L--V--I--VSN.....	152
F_BZ163A	-T-----NAI.....	----A.....	139
G_LBV217	-----E-----NAI.....	A-V--S-K-S-NSSL.....	149
O_ANT70	Q-----QM-F--QME--NIA.....	GT--.....	139
O_MVP5180	-----E-----E-M-F--QM--V-LQ.....	T-K..GLL-E-I--.....	148
O_VAU	-----D-----QM-F--QM---IK.....	SI--T-SPLNS.....	147
CPZGAB	-----Q-SKANFS.....	QAK-L--Q-SSPPL.....	150
CPZANT	T--MXQ-FQ--H-----M-IKM--GY.....	G--P--TPS-T-STVTPK..TT-P.....	145
A_ROD	I--VWH-FET-I-----AMK-SSTESSTG...NNNTTSKS-.S.-TT--P-DQEQE--.S-DTP-A-A	149	
B_EHOA	TK-VW--FET-I-----MK-NKTWSS.....	ASKE-.T..SSASLRSSTQTLL--D-K--QN	148
SD_MM251	I--VWQ-FET-I-----S--I-MR-NKSETDRWG.LTKSSTTI-TAAPTSAPVSE---V-----A-	156	
STM_STM	I--VWN-FET-I-----IAMR-NKNETDKWGLTGKTVTTV-.P.--AAAAA-KPE.LV-----VSN	154	
VER_AGM3	GSN-HL-FE-T-----S--IKMS-VELNSSEPT..TTPKST.-.AS-T-I-AS-TTLPCVQ-K--TVLES	156	
GRI_AGM677	ESN-HL-FE-TMR-----S-I-IKMS-VEL.....	-G-.A..TKA-T-ATTT-T..TPCQN-STE	140
TAN_AGMB14	ESN-HL-FEST-----M-IRM--RLPSPTP...SSSSTTR-.T.R-PCPG--.....	NDE-NSH	108
SAB_SAB1C	ESN-HL-FEST-----S-M-IKM--YRLEGG.....	AATT-.S.PSTS-ARPEVVSV..GFND-V-E-	150
SYK_SYK	L--MSA-FL-ANR-----A-M-IRM--LD.....	SP.A.-STP-TSPPTPPP----W.....	139
V1 ---> <--- V2 Loop ----->			
	^ * ^ * * ^ ^ ^	^ ^ ^ * ^ ^ ^	
most-likely	DNCTGLEK....GEIK..NCSFNITTEIRDKKQKEYALFYKLDVVPI.D.NN....NTS....YRLINCNTSV		216
A_U455R....E-M-----M---L-----V-S---R--I-Q-.NKTD....N-----T		198
B_HXB2RM-----S-S--G-V----F----II--.---D....T-----K-TS----		200
C_UG268A2	.-GKLM-Q-----V-----TAH-----SL.E.G----SNT.....A		205
D_ELI-E....KGM-----V--VLK----QV----R--I----.---DSS.T-STN.....A		201
E_TN243	...I-NIT....D-VR-----M---L-----VH----I-Q-.G.DK....S-E.....		207
F_BZ163A	-GDLKEGP....A-Q-----V-M-VX--Q--VH--XR--I--.S..GR...-G-GD.Y-----T		200
G_LBV217	G-S-VSSI....E-M-----V--SK-----R-----NG-E----S-TT.D-----A-T		210
O_ANT70E....NLM-..K-E--V--V-K--E-KQ----VS-LMEL.NETSSTNKTN-K..M-T-T--STT		196
O_MVP5180-MR-----V--VLT--EQKQ----VS-LSKVN-.S-A..V-GTT...-M-T--STI		200
O_VAUNNT....K-V--Q-D--V--VLK--QE-KQ----VT-L-K-.NATS....-ETM.....STT		201
CPZGAB-M-----V--L---KQV-S--VE--NL.G.-E----NT....I-----TA		198
CPZANTIV....DGM-LQE-N--QS-GFK----MK-I---G-LMKCQ--....E-NC...-Y-WH---TT		200
A_ROD	--S--GE....E-TI...-Q--M-GLE---KQYNETW-SK--CETN.-ST...-QTQ..CYMNH-----		209
B_EHOA	-S-A-IGL....E-MI..D-Q-KM-GLK--ESKQYKDTW--Q-L-CEKGTRS....-E-K..CYIKT---I		208
SD_MM251	N-----Q....EQMI..S-K-TM-GLK---TKEYNETW-ST-L-CEQG.-ST..D-E-R...CYMNH-----		217
STM_STM	N-----E....ESLV..G-K--M-GLK---REYNETW-SS-LICEQNVTG....EE-R...CYMRH-----		214
VER_AGM3	C-E-II--ELNEEPAS...-T-AMAGYV--Q-K-YSVVWNDAEIMCKKG.--.SNRE...CYM-H--D--		219
GRI_AGM677	QIEGEMAE....EPAS...-T-A-AGYQ--V-KNYSMWT-DQEL-CN.N.KTGS.EKG-K..DCYM-H--D--		202
TAN_AGMB14	LVTNSM-F....ENSS..M-T-AMAGYM--Q-KTYNSTW-DAELMCEPESKK....SRG...CYM-H--D--		168
SAB_SAB1C	E....M-----EQAM....AMAGYR--V-KNYSTVWDDQE--CEEGREKS...-ATHTVGCYM-H-----		210
SYK_SYK	..WGDNST....EPRF.....L-GGFK---QYR-F--D-LMKEEG.....-S-Y...-Y-LH-----		194

	* * * ^^ ^ * ^ * ^	
most-likely	ITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGT.GPCKNVSTVQCTHGIKPVVSTQLL.NGSLAEEEIV	287
A_U455	-----K-PE---K.-R-----R-----R-R	269
B_HXB2R	-----N-T-----T-----R-----V-	271
C_UG268A2	-----TLD-----Y-----N-T-----N-----I-----I-----I	276
D_ELI	-----R-----R-----T-----R-----VI	272
E_TN243	-K-----I-D-----T-Y-F-----N-----S-----I	278
F_BZ163A	-----WG-----Y-----X-----D-I	271
G_LBV217	-K-----T-D-----T-R-----M	281
H_VI557		NHVI 4
O_ANT70	-----Y-F---STE---.T-R-ITV-T----R-T----I---T-SKGK-R	267
O_MVP5180	-K-----T-Y-F---TD---.L-H-I-V-T----T----I---T-SR-K-R	271
O_VAU	-R-----C-F---ETG---.L---TV-T----T----I---T-SKGN-T	272
CPZGAB	-----T-----D-S-K.-K-T---H-----T-----I-----GN-T	269
CPZANT	---S-E-ST-----Y---R-E-ED-T-V.-M----V-H---S-M-A-W----T.....	264
A_ROD	--ES-D-HYWDALRFR--P-Y-L-R--TNYS-FAPN-SK-VAST--R.MMETQTSTWFGF--TR--NRTY	281
B_EHOA	-QES-D-HYWDALRFR--P--L-R--T-YS-FMPN-SK-VVS.LYR.MMETQTSTWFGF--TR--NRTY	279
SD_MM251	-QES-D-HYWDALRFR--P-Y-L-R--TNYS-FMPK-SK-VVSS--R.MMETQTSTWFGF--TR--NRTY	289
STM_STM	-QES-D-HYWDALRFR--P-Y-L-R--TNYT-FAPN-SK-VVSS--R.MMETQTSTWFGF--TR--NRTY	286
VER_AGM3	-KE--D-TYWDELRRLR-----L-----YDYA-FKTN-S---V-H--NL-NTT-T-G----YS-NRTQ	291
GRI_AGM677	-KE--D-TYWDTLRLR-----Y-L-----DYR-FAPK----V-H--RL-NTTIT-GIG----RS-NRTE	274
TAN_AGM677	-KE--E-TYWDTLRLR-----AV-----K-TNYT-F.-T-R--V-S--GLMNTT--SAFG----Q--NRTE	239
SAB_SAB1C	-KE--D-TYWDTFRLR-----Y-L-R-A-TDYS-H.KA-R--TVSA--RL-NTT--GIGI---YVANRTE	281
SYK_SYK	-SA--E-QT-Q-F--Q---P-YSL-----TN-E-D.DV-T--TA-S--Q.....FNT	247
<- V3 neutralization loop		
	^ ^ ^ ^ ^ ^	
most-likely	IRSENFTDNAKTIIVQLNES.....VEINCTRPNN..NTRKSITI..GP	327
A_U455	-----N-----VNP-----K---S---Y-TRK-I..RYS---S	311
B_HXB2R	---V-----T-----R-R-QR--	313
C_UG268A2	-----L-N---I-----K-----A-----E--R----	316
D_ELI	-----L-N---N--AH-----K-T-A--YQ----QRTP---L	312
E_TN243	-----L-----H-K-----G-----S-----XP---X---	318
F_BZ163A	--Q-IS-----HF-----Q-----G-H----	311
G_LBV217	-----N---N---F-K-----ID-V-----H----	321
H_VI557	---K-I---T-N-----KSP-----P-----S----	44
O_ANT70	MMAKDILEGG-N---T--ST.....LNMT-E--QI...DIQ.EMR----.	305
O_MVP5180	-MGK-I-ES--N---T--TP.....INMT-I-EGI...AEVQD-YT----	311
O_VAU	-MGK-IS-SGEN-LIT--TN.....IT-A-E--G...Q-IQK-MA...	311
CPZGAB	V-V--KSK-TDVW---V-A.....SL--H--G...---GEVQ----	309
CPZANTY-T-T-VVMNGRKNESV.....LVRFGKEFENLT-T-I--G...R-VRNLQ----	311
A_ROD	-YWHGR---R---S--KYY.....NLSLH-K--G...KIV-Q-ML..MS	320
B_EHOA	-YWHGK---R---S--SYY.....NLTMH-K--G...KMVVP-RT..VS	318
SD_MM251	-YWHGK---R---S--KYY.....NLTMK-R--G...K-VLPV--..MS	328
STM_STM	-YWHGK---R---S--KYY.....NLTMS-R--G...K-VLPV--..MS	325
VER_AGM3	-WQKHRVS-.DSVL-LF-KHY.....NLTVT-K--G...K-VLPV--..MA	332
GRI_AGM677	-WQKGGN--.D-V-IK--KFY.....NLTVR-R--G...K-VLPV--..MA	315
TAN_AGM677	-WQKHGV-S-DSV--K--KHY.....NLT-L-R--G...K-VLPV--..MA	280
SAB_SAB1C	-WQK-GNS-DSV--R--RYF.....NLT-R-R--G...K-VLPV--..MA	322
SYK_SYK	LA-TW-.....GTYKAKDKVRFIKQDKKNESVIIILVPEALRLQ-I-E--G...ESI-N-QLAA..	307

HIV1 ENV

	V3 neutralization loop ->	
	* ^ ^ ^ ^ ^ ^	
most-likely	GQAFYAT...GD.I.....IG.DIRQAHCNISGAKWNETLQQVA...KK.....L...REQFG.	369
A_U455	----V---K.-.....--V-RRD--R-I----.EQ.....-KKK-N.	353
B_HXB2R	-R-.FV...IG.K.....NM-----R---N-K-I...S-----N	355
C_UG268A2	--T-----Y---RNE--I---W-R..E-----KRH-P.	358
D_ELI	--SL-T-R..SRS-----R-Q-SK-----R-----GTLL..	352
E_TN243	--V--R-----XR-Y-E-N-T---RV-K--T..E-----K-H-N.	360
F_BZ163A	-R-----K---V--TQ--K--E--R..A-----KSH-P.	353
G_LBV217	--L-----A-----V-ETD-RDM-K-K..AQ-----QGIY..	362
H_VI557	-----K-Y---TRED-KR-HE-V..QQ-----H-N.	86
O_ANT70	PM-W-SMG.I-G.T....A-NSS-A-Y-KYNATD-GKI-K-T-ERYLE.	349
O_MVP5180	MRWRSM-LKRSN.N.....TSPRS-V-Y-TYNKTV-ENA---T-IRYLN.	360
O_VAU	PM-W-SMA.LSN.T...K--T-A-Y--Y-ATD--KA-KNITERYLE.	359
CPZGAB	-MT--NI...EN.V.....V--T-S-Y-K-N-TT--R-VEE-K...-A.....ATSS..	350
CPZANT	-MT--.N..VE.-A....T--T-K-F-TVNKTL-EQARNKTE ..HV.....A-HWKK	354
A_ROD	-HV-HSH..YQ.P....-NKRPA--W-WFK--KDAM-E-KETLA-.....HP...-YRGRT.	366
B_EHOA	-IL.F.H...SQ.P....-NKRPK--W-WFK-.N-T-AI-E-KETI-N.....HP...-YSGTT	363
SD_MM251	-LV.F.H...SQ.P....-NDRPK--W-WFG--KDAIKE-KQTIV-.....HP...-YTGT.	372
STM_STM	-LV.F.H...SQ.P....-NERPK--W-WFG-.E-RGAIKE-KETLV-.....HP...-YTGT.	369
VER_AGM3	-LV.F.H...SQ.R.....YNTRL--W-HFQ-.N-RGAWKE-KNEIV-.....-PKD-Y-GT.	378
GRI_AGM677	-LV.F.H...SQ.K.....YNTRLK--W-HFQ-.D-KGAWKE-REEV--VKNLTEVSIENIH-...-RIW..	370
TAN_AGMB14	-MV.F.H...SQ.K.....YNTKL--W-HFQ-.D-KGAWKE-RETIV-.....-PKHKY-GT.	326
SAB_SAB1C	-LV.F.H...SQ.K.....YNTRLK--W-WFG-.N-RGAWKE-KETIVR.....-PPKKYSGT.	368
SYK_SYKF#...SQ.SYK#KLKT-R-AKR-F-RVT-.N-T-FFK--H..EQ.....A...TKTWK.	348
	<-- V4 Loop	
	CD4 CD4	
	^ ^ ^ * * ^ ^ ^ ^ ^ ^ ^ ^ ^	
most-likely	.NKT....IIFNQSS.GGDPEITTHSFNCGGEFFYCNTTQLFN.STWN...NGTW.....NS..	416
A_U455-AS---I-----SG---.I---SMS.....-D..	400
B_HXB2R-K-----V-----S-----F...-S--STE.....GS-N..	407
C_UG268A2-N-T-P---L-----R-----SS---.SD..N-S-----	403
D_ELII...K-KP---SG-----ISAW.....-NI.	401
E_TN243-QPP---L--M-H--R-----R--,N-CI..G-E-M.....	407
F_BZ163A	..-A...-K-S---L--M---R-----SG--D.....	392
G_LBV217	..-S...-T-S---L--A---R-----SG---.SIL..S-NNAP.....S..	410
H_VI557	..-Q---EP---M---M-T--R-----SK---.V...-S-S.....-DI.	134
O_ANT70	..-N-GSINMT--H---L-V-HLH--H-----AKM--Y-FS..C---TCS.....VSN..	404
O_MVP5180	..-V....-SRT...-A-VSHLH--H-----SGM...Y-FI...-C-KSGCQE...IKGS...	415
O_VAU	..DV....MK-GNH...-E-A-V-NFF--H-----NR--.H-FSCKK-M-NNKI.....NCT-IS.	416
CPZGAB	..-R-AA.N-TL-RA...-V-H-M-----S-I-TD.....-I.....	394
CPZANT	VD---NAKT-WTF-D...-VKV-W--Q-----DI-PW--.A-YT..G-LI.....	403
A_ROD	..-D-R..N-S-AAPGK-S---VAYMWT--R--L--M-WFL..WIE.....-K.....	412
B_EHOA	..-ISQ...-RLAEHARSS--VRYMWT--R--L--M-FFL..WVE.....-R.....	408
SD_MM251	..-N-D..K-NLTAPG.----V-FMWT--R--L--KMNWFL..WVED..RDV-TQR.....PK..	424
STM_STM	..-D-A..K-RIVAPG.----V-FMWT--R--L--KMNWFL..WIE...RS-SEMRD.....WNK-K..	425
VER_AGM3	..-D-E...E-YL-RL.F---AANLW--Q-----KMDWFL..YLN....R-VDPDH.....NPC-G..	432
GRI_AGM677SANFW--Q-----KMDWFI..YLN....R-ED.....AEG..	407
TAN_AGMB14	..-D-K..K-WLRRQ..W---AASIW--L-----TPDWFL..YLN....-ESSEGSFTDVEGNRCS-ITS	388
SAB_SAB1C	..-D-N...K--L-RQ.W---SEFFF--Q-----KMDWFL..YLN....-KSVDPDHN...NCAK-N..	424
SYK_SYK	..-V....NTTWR-QP---L-VR--W-Q-----VSK--A.NIT.....-NAS.....K-N..	397

	V4 Loop	-->	
			CD4
	^ ^ ^	*	* ^ ^ ^
most-likely	...TES.....	NDTITLPCRIKQIINMWQEVGKAMYAPPIEGQITCSSNITGLLTRDGGDN....	470
A_U455	...MGP.....	-G---Q-----R-Q-----Q-V-R-E-----T....	453
B_HXB2R	...-G.....	S-----K-----S-----R-----S... .	461
C_UG268A2	I-----G-----K-----R-----K-K-----R-----ET....	452
D_ELI-NN.....	ST-TN--Q-----K-XWQA-X-I-----RN-L-----I....	459
E_TN243GC.....	G-----K-----GA-Q-----S-K-N-V-----I-----A....	460
F_BZ163AG.....	-G---L----MV----R---A-SEN---N-----Q-N....	446
G_LBV217N.....	SKN-----VR--R-Q-----A-N-----K-----I-----S... .	463
H_VI557SNG.....	TTI-E-----R-Q-----K-K-----E-----I-----S-HV.-S... .	188
O_ANT70VSQ.....	GNNG---KLR-VVRS-IRGQSGL---K-NL---M-----MI-QM-NT.W....	457
O_MVP5180-T.....	-KNG-I---KLR-LVRS-MKGESRI---P-NL---H-----MI-QL-QP.W....	467
O_VAUNN.....	G-QAI---LR-VVRD-MRG-SGL---P-NLV-R-----MI-QL-TPW-KTHP	474
CPZGAB	-GI-I-----R-VSS-MR--RGI-----R-N---N-----S-TPVT....	445
CPZANTT	-GALIAH-----V-H-GI-S-GI-LA-RR-NVS-T-S---IM-E...-QIY...	453
A_ROD	HRNYA-H-----T-HK---RNV-L---R--ELS-N-TV-SIIANI-WQN....	462
B_EHOAG.....	LKRNYAS-H-R--V-T-HKI-RNV-L---R--ELS-N-TV-S-IANI-WI-K....	460
SD_MM251-R.....	HRRNYV--H-R---T-HK---NV-L---R--DL---N-TV-S-IANI-WT-G....	477
STM_STMK-Q.....	QKRNYV--H-R-V--T-HK---NV-L---RQ-DL---N-TV-SIIANI-WTN....	479
VER_AGM3-KGKGKA	PGPCAQRRTYVA-H-RSV--D-YTLSRKT---R---HLQ-T-TV--MSVELNYNSK....	494
GRI_AGM677-NRTC	DKGKPGPGPCVQRRTYVA-H-R-VV-D-YT-S-KV---R--HLE-N-SV-A-YVAI-YN....	472
TAN_AGM677	GGL-G-TR.....	KCLKRTYVGLH-RSVD-D-YTLLQKR---R--HLE-R-TV-SM-VSL-YN-K....	449
SAB_SAB1C	...-KPC.....	WQRTYV--H-R-VV-D-YTLS-KT---R--HLE-N-TA-A-YVELNYNSK....	479
SYK_SYK	...YA.....	SNLR-S-A-R---D-RY-R-LI-L---TA-H-K-T--V-AV-TDIEYYPG....	449
		-< fusion peptide->	
	^ ^ ^ ^ ^	gp120 / gp41	
most-likely	NSTN..ETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVVQ....REKRAV..GLGAVFLGFLGA	534	
A_U455	-N-K.N-----K-----R-----E-----I-----	518	
B_HXB2R	-NES..-I-----	525	
C_UG268A2	SE--ST-----EVK-----E-----I-----	518	
D_ELI	-----Q-----R-----E-----I-----M-----	523	
E_TN243	TT----NIK-----Q-----I-----R-----E-----I-----MIF-----	524	
F_BZ163A	-Q-----K-Y---V-----E-----R---Q---R-----M---L-----S	509	
G_LBV217	TNG-----A-----I-K-----R-----G-----I...T-----	526	
H_VI557	AES..V-----RV-----L---R-----E-----E-----	252	
O_ANT70	--S-NNV---I---K-I---T-FN---RVK-FS---RIA-P-ISTR.TH-----ML---V-S-	527	
O_MVP5180	--G.EN-L--V---K-I---TK---N---Q-K-FS---MS-PIINIHTPH-----ML---V-S-	537	
O_VAU	-----L-----K-I---TQ-F---RVK-FS---IA-PTIGTR.SH-----A---AML---I-S-	540	
CPZGAB	-NSG.NL---T---N-K-I-----R-----S-----R-HT-ARQ.KD-Q---AF-----L-----	515	
CPZANT	-E...-VKVS-.AARVA-Q---A-SR-Q---E-X---S---TX---PEIKQH..S-Q---GI..-I-LF---L-S-	518	
A_ROD	-Q---I---SA.EVAE.LY.-L---GD---L-E-T-I-F---E---YSSAH...G-HT-G-F.V---.AT	524	
B_EHOA	-L---I-VSA.EVAE.LY.KL---GD---L-E-T-I-F---SI---YSSVT...P-N---G-L.V---.AT	522	
SD_MM251	-Q-S..I-MSA.EVAE.LY.-L---GD---L-E-T-I-L---DV---YTTGGT..S-N---G-F.V---.AT	540	
STM_STM	-E...-I-ASA.EVAE.LY.-L---GD---L-E-T-I-L---NV---YTTST...S-T---G-F.V---.AT	541	
VER_AGM3	-R...-V-LS-...QIETI-AA---GR---L-E-T-I-F---EVR-YTGHH...D-T---VPF.V---.	556	
GRI_AGM677	K-GPINV-LS-...QV-SI-AY---GD---L-E-T-I-F---DVR-YTGPT....VPF.V---.	535	
TAN_AGM677	-M---V-LTA...NLENI-AY---GR---LIE-K-I-F---EVR	487	
SAB_SAB1C	-R...-V-LS-...QIESI-AN---GD---L-E-K-I-F---VR-YTGPE....Q---VPF.V---.	540	
SYK_SYK	STL---T-TANVE.-V---AD-FN---LIQ-K-I-F---DQR-YELPN...T---APLA---L-S-	512	

HIV1 ENV

	<-- immunodominant	
	*	
most-likely	AGSTMGAASITLTQARQLLSGIVQQQNNLLRAIFAAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGC	607
A_U455	-----S-----K-----Q-----	591
B_HXB2R	-----M-----I-----	598
C_UG268A2	-----M-----T-----I-----Q-----	591
D_ELI	-----R-V-----M-----I-----	596
E_TN243	-----XS-----X-----XX-----X-----K-X-L-X-----	597
F_BZ163A	-----S-----Q-----L-----	582
G_LBV217	-----LA-----V-----Q-----	599
H_VI557	-----V-----S-----Q-----M-K-----	296
O_ANT70	-----AT-----A-----THT-----K-----D-----Q-----Q-----R-SX-----R-----R-----L-----L-----TL-----QN-----SL-----	600
O_MVP5180	-----ATA-----RTHSV-----K-----D-----Q-----R-----S-----R-----R-----LQ-----L-----TLIQN-----R-----NL-----	610
O_VAU	-----ATA-----RTQH-----IK-----D-----Q-----RPS-----R-----R-----L-----L-----TFIQN-----NL-----	613
CPZGAB	-----AV-----K-----SI-----V-----L-----Q-----I-----L-----	588
CPZANT	-----A-----A-----T-----N-----XH-----A-----Q-----T-----S-----V-----M-----K-----R-----SL-----	591
A_ROD	---A-----L-----VSA-----S-----T-----A-----QQ-----DVVKR-----E-----R-----T-----N-----T-----I-----K-----Q-----AR-----NS-----	597
B_EHOA	---A-----L-----SA-----S-----T-----A-----QQ-----VDVVKR-----E-----R-----T-----N-----T-----I-----K-----AQ-----NS-----	595
C_2238	DVVKR-----EM-----R-----T-----N-----T-----T-----I-----K-----AK-----NS-----	42
D_F0784	DVVKR-----E-----R-----T-----N-----T-----T-----I-----K-----AQ-----NS-----	42
SD_MM251	---A-----L-----A-----S-----T-----A-----QQ-----DVVKR-----E-----R-----T-----N-----T-----T-----I-----K-----AQ-----NA-----	613
STM_STM	---A-----L-----A-----S-----T-----T-----QQ-----DVVKR-----E-----R-----T-----N-----T-----T-----I-----K-----AQ-----NS-----	614
U_SMCI2	DVVTR-----E-----R-----T-----N-----T-----T-----I-----K-----AR-----NA-----	42
VER_AGM3	--TA-----ATA-----SQH-----A-----L-----K-----A-----V-----QM-----K-----I-----V-----N-----N-----T-----L-----K-----E-----AR-----NA-----	629
GRI_AGM677	--TA-----AT-----S-----H-----A-----L-----K-----A-----V-----Q-----Q-----K-----I-----V-----N-----N-----T-----L-----K-----E-----AR-----NS-----	608
SAB_SAB1C	--AA-----ATA-----SQ-----A-----L-----K-----A-----V-----Q-----QM-----K-----I-----V-----N-----N-----T-----L-----K-----E-----AR-----N-----	613
SYK_SYK	--TA-----G-----ATA-----L-----SQT-----A-----QK-----E-----V-----G-----V-----N-----N-----LT-----L-----T-----R-----AI-----SN-----	585
region ->		
	* ^ ^ ^ ^ ^ ^ ^ ^	
most-likely	SGKLICTTTVPW..NSSW..SNKSLT.....PIWNNTMWEWEREIDNYTALIYTLLEESQNQQEKNEQ	667
A_U455	-----QE.....D-----LQ-----K-----SS-----GI-----Q-----I-----L	651
B_HXB2R	-----A-----A-----E.....Q-----HT-----D-----N-----S-----HS-----I-----	658
C_UG268A2	-----A-----G.....D-----D-----Q-----D-----S-----GT-----R-----D-----K	651
D_ELI	---H-----N-----R-----N.....E-----Q-----G-----S-----I-----T-----K	656
E_TN243	---I-----A-----T-----R-----FE.....E-----X-----I-----X-----S-----NQ-----EI-----T-----DR-----K	657
F_BZ163A	-----N-----QE.....E-----E-----QK-----S-----SNEV-----R-----I-----K-----	642
G_LBV217	-----N-----T-----FN.....E-----D-----I-----N-----HQ-----S-----I-----	659
O_ANT70	K-----V-----Y-----S-----K-----RT-----I-----G-----E-----DTL-----Q-----D-----Q-----S-----ISST-----EEIQKA-----V-----Q-----K	658
O_MVP5180	K-----Y-----S-----K-----T-----GRYNDD.....S-----D-----L-----QQ-----DQH-----N-----VSSI-----DEIQAA-----D-----VK	671
O_VAU	KNR-----Y-----S-----K-----KT-----GGD-----E-----DEL-----QQ-----DQH-----N-----VSSF-----EKIQ-----A-----E-----K	672
CPZGAB	---AV-----Y-----N-----PG-----ST-----D.....D-----G-----L-----QQ-----DKLVS-----GK-----FG-----A-----S-----R	649
CPZANT	AD-----VT-----H-----N-----V-----FTQ-----CAKNSSDIQC-----E-----Q-----D-----LVQ-----S-----GQ-----NI-----QIAHE-----R-----KK	660
A_ROD	AFRQV-----H-----V-----D-----A.....D-----D-----Q-----KQVRYLE-----N-----SKS-----QA-----I-----MY	653
B_EHOA	AFRQV-----H-----V-----E-----K.....D-----QQ-----QVRFLD-----N-----TK-----A-----I-----MY	651
C_2238	AFRQV-----H-----V-----D-----MQ.....E-----Q-----QQ-----KQ-----AFLEDN-----TE-----QA-----I-----MY	98
D_F0784	AFRQV-----H-----E-----P-----T.....D-----D-----Q-----QV-----FLE-----N-----TQ-----A-----I-----MY	98
SD_MM251	AFRQV-----H-----P-----A.....D-----D-----D-----Q-----KV-----FLEEN-----TA-----A-----I-----MY	669
STM_STM	AFRQV-----H-----P-----D.....V.....D-----Q-----KV-----FLE-----N-----TQ-----A-----V-----MY	670
U_SMCI2	AFRQV-----H-----E-----NT.....Q.....D-----Q-----IK-----RDLE-----N-----SES-----A-----I-----MY	98
VER_AGM3	AW-----QV-----H-----QW-----NR.....D-----L-----Q-----SYLEGN-----T-----Q-----ARA-----E-----LD	685
GRI_AGM677	AW-----QV-----H-----KY-----NT.....K-----D-----L-----Q-----NALEGN-----TQ-----A-----ES-----LD	663
SAB_SAB1C	AFRQV-----H-----L-----KY-----NT.....D-----E-----Q-----Q-----EK-----E-----N-----SRI-----QAHE-----EQ-----LD	668
SYK_SYK	AF-----Q-----H-----A-----T-----E-----KACGN-----N-----FCPK.....Q-----K-----HR-----Q-----V-----L-----DH-----DG-----R-----A-----E-----R-----VH	648

		\ / 3' sj	
most-likely	ELLELDKWASLWNWFDITNWLYIKIFIMIVGGGLIGLRLIVFAVLSIVNVRQGYSP.LSFPPGYIQQTHLPAP		739
A_U455	D--A----N-X---N-S----RL-VI-----T---I-----.	--LA-I-	717
B_HXB2R	-----N-----L-----V-----.	-----T-	724
C_UG268A2	D--A----QN--S----K-----I-----K-----.	--LT-T-	717
D_ELI	-----S--Q-----I-----L-----.	--L---	722
E_TN243	D-----S-----I--X-----.	--L.....--PTHHQ	723
F_BZ163A	G--A-----S-S-----.	KA-----.	708
G_LBV217	D--A----E--S--S--Q-----K-----.	--LAHHQ	725
O_ANT70	K----E---I---L---K----A-I---A-V-V-VIMI--N-KNI---Q---L-----IPNNHQ	724	
O_MVP5180	A----E-----K-----A-I---A---I-VIMII-NL-KNI---Q---L-----IPV-HR	737	
O_VAU	-----E---I---L---K----A-I---A---V-V-MI---NL-KNI---Q---L-----IPIQQQ	738	
CPZGAB	D-----Q-----K-----L-A---I-----IMT-F-V-R-----.	--LI-VQ	715
CPZANT	--Y-----S-----Q-----AIV-----LLVLV-CLRK-----H-----.	-IPTQNZ	726
A_ROD	--QK-NS-DIFG---L-S-VK--QYGVL-IVAV-A---IY-VQMLS-L-K---R-VF-S-----I-IHKD	726	
B_EHOA	--QK-NQ-DIFS---F-S-MA--RLGLY--I-IVV---AIYIIQMLA-L-K---R-VF-S---S-T--IPIRKD	724	
C_2238	--QK-NS-DVFG--L-L-S-VK--YLGFY--A-V-V---AIY-VQMLM-L-K	151	
D_F0784	--QK-NN-DIFG---L-S-IK--QYGVF--V-I-L---IIY-VQMLAKLSK	150	
SD_MM251	--QK-NS-DVFG---LAS-IK--QYG-YV-V-L---IYIVQMLAKL---R-VF-S---S-F---TQOD	741	
STM_STM	--QK-NS-DVFG---L-S-VR--QYGVYL-I---VM--VAIYIMQLLA-L-K---R-VF-S---SCR--IPIHKG	743	
U_SMCI2	--QK-NS-DVFG-R---F-S-VK--Y-GFYV-V-I---IVYLIQLLGKL-K	150	
VER_AGM3	AYQK-SS-SDF-S---FSK--NIL--GFLD-L-I---LLYT-Y-CIA-----.	--P.....-I-IHPW	751
GRI_AGM677	LYQK--D-SGF-S---SLST--G-V--GFLVIVIIL--FAWVLWGCIERNI---N..-P.....-I-IHSS	728	
SAB_SAB1C	SYQK-VS-SDF-S---L-K-FGWM--A--VIA-I-VA-VLLVIIG-LRKF-K--A--SL-SSH\$	732	
SYK_SYK	D-TK-QE-D---S---LSK-FF--L-GFYVI-A-VL--L-SFSVG-IKNLLG--V-I-QN.....P-QGRKD	715	
<- tat cds			
most-likely	RG.PDRPEGIEEGGERDRDRS....WRL.VN.G.....	FLALIWDDLRSLCLFSYHRLRD	788
A_U455	E-.LG--GR-----QGK-----I---S.-.	IA-----N-----	766
B_HXB2R	--.-----I-----.	S-----	773
C_UG268A2	--.-----GE-----Q-----V--.IS.-.	A-----R-----	766
D_ELI	--.-----T-----G-----V--.L--.	S-----	771
E_TN243	-E.L---R--G---QG-E-----V--.S.-.	A-----	772
F_BZ163A	-E.-----E--G---QG-----V--.D.-.	V-----N-----H--	757
G_LBV217	-E.-----R--G---Q-----V--.S.-.	SIA-----Q--	774
O_ANT70	EE.AGT-GRTGGG--EG-P-W....IPS.PQ.-.	P-LYT--TII-WT--L-SN	773
O_MVP5180	QE.AET-GRTG----G--PKW....TA-.PP.-.	QQLYT--TII-WT--L-SN	786
O_VAU	AE.VGT-G-TG-G--DE--R-W....TP-.PQ.-.	H-LYT--TII-WI--L-SN	787
CPZGAB	-E.QG-LGE-D-G--Q--S-----V--.E.-.	C-P-----N-GIW--QS-TS	764
CPZANT	QD.-EQ--E-R---RK--I-W....RA.QH.-.	F--L	758
A_ROD	--Q-ANE-.T--D--SNGG--Y....PWPIAYI.	HFLI.RQLI-L-T.RL-SIC--	775
B_EHOA	--Q-ANE-.T--G--NNEGY----PWQIEYI.	HFPI..RQ--D-LIWL-SGC-T	773
SD_MM251	PAL-T-EG.K-GD---GGGNS----PWQIEYI.	HFLI.RQLI-L-T.WLFNSNC-T	790
STM_STM	QEQTKEG.T--G--D-GGIN----PWQIEYI.	HFLI.RQLV-L-T.WL-NNF-A	792
VER_AGM3	K-Q--NA-.GPG---DKRKS----EPWQKES-TAEWKSNWCKRLTNWCISSSI.....	WL-NSCLT	809
GRI_AGM677	AER--NGG....-QD-GGES-SSKLI--.QE.E.....	S-INN	764
SYK_SYK	P-K-ADE----GSG--E....GLN.-S.T.....	SRE.....	740

HIV1 ENV

most-likely	LL....L.IV....ARIVELLGRSSLKGLRRGWEALKYLWNLLQY.....	WS.	826
A_U455	FA.....-.-.-A-----L--G-----L-.....	-G.	804
B_HXB2R	--.-.-.-T-----W-----	--.	804
C_UG268A2	--.-.-A...-A-----R-Q-----GS-V-----	-G.	804
D_ELI	-I.....-A....V-----DI-----	-----	802
E_TN243	FI....S.-A....-A----H-----G---G---L-----	-G.	810
F_BZ163A	--.-.-T....-.-NR.....-----L-G--TL-----	-G.	788
G_LBV217	FI.....T-T----N-----L---G-----L-----	-G.	812
O_ANT70	-A....SG-QK...VI SYLR--LWI..LGQKII INV CRICA AVT--	-----L.	811
O_MVP5180	-I....SG-RR...LID YLG--LWI..LGQKTI--CRLCGAVM-----	-----L.	824
O_VAU	-A....SE-QK...LIRHLG--LWI..IGQ-TI--CRLFKAII-----	-----L.	825
CPZGAB	-A....C.N----W-QLKT--HLI-HS--LLR-R-CL-GGI-----	-G.	802
A_ROD	-----SRSF....LTLQLIYQN.....L-D-LR-RT..AF--GC.....	E-I.	809
B_EHOA	-----SKTF....QTLQPV-QP-----RLPPAY-R-GI-----	S-F.	803
SD_MM251	-----SRAY....QILQPI-Q-L-A.T--VR-V-RTELTY--G-----	-----Y	830
STM_STM	C.....SR-Y....QTLHPTFQ-I-R.I-Q-IR-VVRLGAAY--GC-----	I-I.	833
VER_AGM3	--VHLRS.AF....QY-QYG--ELKA.AAQ....VVA-AR-A-N.....	AGY	848
GRI_AGM677	WW....-NFKSCSL--RTWCYNIC-TL-IFIRT-VG-----	GL.	802
SYK_SYK	S-R-SLEAGQQLWRTVCSSFRSLIRQLTIT-GF	773

-< rev cds
 ^ ^ ^

most-likelyQELKNSAVSLLNATAIAVAEGTDRVIEVLQRIGRAILHIPRRIQGLERALL		878
A_U455R---I--IT--D-V-V---GWI----IG-T----N-----		856
B_HXB2R-----V-GAC---R-----I-----		856
C_UG268A2-K--I---DTI---S----I---G-G-----F-A--Q		856
D_ELI-R--S--FD-I-----II--AC--V-N-----S--		854
E_TN243-I--I-----GW--K---A-GAW---		847
F_BZ163A-IG-F-T--V-----I-AA-----N--T-----F-----		840
G_LBV217-IN--DTV--T-NW-----A--AC---N-----		864
O_ANT70-Q--T---DTL-V--NW--GI-AGI-----TG-RN-----S--		863
O_MVP5180-TN--DTI--VS--NW--GI-LG-----QGF-----A--I-V		876
O_VAU-QT--TN--DTV-V---NW--ST-LGI-S---G--N-----L--		877
CPZGABK--I--I--D-----I--AF-VTL-I-RN-----		854
A_ROD-AFQ-AARATRETL-G-CR.GL.WR--E---G--AV-----A-I---		858
B_EHOA-AIQ-AARAAGETL-S-AR.TS.WG--R-AAGE-IA-----A-L---		852
SD_MM251FH-.AVQ-GWRSATETL-G-WR--L.W-T-R-G--W--A-----LT--		880
STM_STMAAQ-AWRAAGETL-S-GR.--L.W-T-G-V--R-GA-----LT--		882
VER_AGM3-I.....WLACRSAY--INS--V-----GI-N		877
GRI_AGM677-Q-QEA-TG-AQ-L-R-AR-AWG-LGAI VRSAY--VINS--V-----KV-G		854
SYK_SYK	ISYGFN---IA-A--GREIRDW--AIWQAIYAATR-VVE-VAAL---L----IY-N		830

COBBLER sequence from MOTIF

```
>Env C_UG268A2, with embedded consensus blocks
mrvmgijqrncqqwwiwiglfgwilmicnvmgnlWVTVYYGVPVWKDATTTLCATdakay
etevhnvwathacvpptdpnpqevlenvtesfnmWKNTMVEQMHEIDIISLWEQSLKPCVK
LTPLCIKMNCNTvnvnvnitnnanatnspyengkImegqeikncsfnvttieirdkkqtahal
fykldvsvslegnsntyrlighCNTSVITQACPKVYWDPIPIHYCAPAGYAILKCNDTNFNG
tgpcnnvstvqcthgikpvistqllngsiaeiiiirsenltnnakiiivqlnksvein
carpnntntresirigpgqftyatgdigdirqaycnrnewnitlqwvreklkrhfpnk
tinftqpsGGDPEVTHWFNCRGEFFYCNTSWFFNssdnnnstiitLPCRIRQIVNDWHK
VGKAIYAPPIEGNITCNSNITGLLLTLdgetsetnstatfrpgGGDMKDIWRSELYKYK
VVEIKPIGVAPTKVKrvvrekravgigavflGFLGAAGSTMGAASTTLTVQSRQLLSG
IVQQQQNLLRAIEAQQHLLRLTVWGVVKQ1QARVLAIEKYLQDQQLLNWGCAGKLICHTT
VPWnsswsnksgldiwdnmTWQEWERQINNYTGNIYQLEEAQNQQEKNEKLLKLDKWS
SFWNWFDITKWLWYikifimivggliglriifavlsivnrvrkgysplsfqtltptprgp
drpgeieeggeqdrrdrsvrlisgfalawddrlsclfysrrldllliaaravellgr
sslrlglqrgwealkylgslvqywqgqelksaislldtiaiavsegtdriievqgqgigrai
lhiprrirqfleaalq
```

	<- vpU cds signal peptide / gp120	*
CONSENSUS-A	Mrvmgiq?nyq?l.wr??....??W.gtmilg??iIc.na??e.?lWVtVyYGVPVWkdaeT..TLfc	47
CONSENSUS-B	?---k---rk---h----????---l---mlm--s---------e-t----	51
CONSENSUS-C	---?--?r-?gw-i...??---ilgfwmhm--v-g.n-----e-k?---	49
CONSENSUS-D	--r?-er---h----???---L--mLM--sv.a?-----E-t----	50
CONSENSUS-E	--Ket-m-wpn-.k----l---lv--?s-.Sd.N-----r-d----	53
CONSENSUS-F	--R-M-R-W-H-.GK.....LLF--iL----n-----e-T----	52
CONSENSUS-G	-?k---r-W-H-.k----L---LV--.s-.sn.n-----E-D----	52
CONSENSUS-H		..
CONSENSUS-O	-?tMk?MeKrN?.Kl.....-il?l?mALi-P-.LS.-n?q-YA--?----E--?P..v--	46
CONSENSUS-U	---e--R-cp?-.-?.....-il--m-m---.?ns-----	47
CONSENSUS-CPZ	-??????-??.?--?....?????-?--?T.--.?-?----?--?P..?--	18
	* ^^^	*
CONSENSUS-A	ASdAkAydtE?HNW?aTHaCVPTDPnPqEi?le.NVTE?FnmwkNnMVeQmheDiiSLWD.qSLkPCvk	111
CONSENSUS-B	-----v-----vv-?----n-----	117
CONSENSUS-C	-----e-v-----?v----n----d-d-----	115
CONSENSUS-D	---s-k?-a-i-.----N-----	115
CONSENSUS-E	----He-v-----h----n-----q-v----?	120
CONSENSUS-F	----S-Ek-v-----Vv-----n-d-----T-----	119
CONSENSUS-G	----s-s-----n-----E-----	118
CONSENSUS-O	---NLTS--q--I--sQ-----t-?y-p-k---dn-I--Y--?-----?-----q	109
CONSENSUS-U	----s-?-----s---?----n-----?	108
CONSENSUS-CPZ	-??-?S-----?----?----?V--?----?----?----?----?----?----?	54
	* ^ * ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^	
CONSENSUS-A	LtPLcvtl?c.?????????n?t?????????n?t?????????n?????....?????????.....	125
CONSENSUS-B	-----n.td-----?----?-----?----?-----?----?....???	132
CONSENSUS-C	-----n.-----?-----t----?....-----	131
CONSENSUS-D	-----n.t-----?-----?-----?-----?....-----	130
CONSENSUS-E	-----n.tna-----l-----nv--i-nvsniig-it.....	151
CONSENSUS-F	-----?a-----?-----?-----q-----?tl	134
CONSENSUS-G	-----n.-----v-t-----?-----NcT--?n--nNstv-...?????	141
CONSENSUS-O	M-F---QMn.t-----i-----?-----?-----?-----	125
CONSENSUS-U	-----n.t-----?-----?-----	122
CONSENSUS-CPZ	---?----?-----?-----?-----P?????..?	60
	* ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^	
CONSENSUS-A	.m?..?e.....ikncsfnmttelrdkkqkvyslfyrlldvvqi?????????n?????....n???????	160
CONSENSUS-B	?-ek?g-?????-----i--si---v--e-a---k---p-d-----?-----?-----	170
CONSENSUS-C	?-?----?-----?-----a-----i-pl-----?-----s-----	164
CONSENSUS-D	?-?..g.....m-----i-?v---kq-ha---k-----?-----t----	164
CONSENSUS-E	?-..d-----Vr-----hA---k---i-----s-----?-----?	186
CONSENSUS-F	keep.ga.....-Q-----v--Ql--Ha-----I-p-s-----ns-----?-----	175
CONSENSUS-G	?-?..e-?..?m-----i---i---ktE-A---k---p-n---?ss-----?-----	177
CONSENSUS-O	?-?..n-?...?-----q-e---v--V-k---E-KQA---Vs-L-k?n-ts-----T-----	161
CONSENSUS-U	...?-?...?-----?-----?-----kta?-----k---p-nd-----?-----	152
CONSENSUS-CPZ	?-?..?...?-----?-----?-----?-----?-----?-----?-----T.....	73
	* ^ ^ ^ * ^ ^ ^ * ^ ^ ^ * ^ ^ ^ *	
CONSENSUS-A	?yrlincntsaitqacpkvsfepipihycapagfailkc?dk?fnngpcknvstvqcthg?ikpvvst	225
CONSENSUS-B	-----s----v-----n-k-----t-----?-----r-----	237
CONSENSUS-C	-----d-----y-----nn-t-----?-----	230
CONSENSUS-D	-----t-----n-k-----?-----r-----	231
CONSENSUS-E	-----V-k-----i-D-----t-----y-----N-n-----S-----	253
CONSENSUS-F	-----T-----Wd-----Y-----N-k-----	242
CONSENSUS-G	sd-----v-t-K-----?-----d-----r-----	244
CONSENSUS-O	.m-t?-----STt-k-----y-----F---N?Te-----?-----itV-T-----T-----	226
CONSENSUS-U	.-----?-----n-K-----	218
CONSENSUS-CPZ	---?----?-----T?-----?-----?-----?-----?-----?D-?-----?-----H-----?-----?	119
	^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^	<- V3 neutralization loop
CONSENSUS-A	qlll.ngslae??ev?irSenitnNaktiiVql??pV?InCtRp.nnntr.ks??vrif??gpGq??af	278
CONSENSUS-B	-----e---v---f-d-----nes-e-----?-----ih-----r...--	295
CONSENSUS-C	-----i-----l---?-----h-n-s-e-v-----i-----t-----	285
CONSENSUS-D	-----E.-iI-----l---?-----nes-----?y-----qr...tp...?---l	285
CONSENSUS-E	-----e.-iI-----L-----h-NKs-e-----s-----t...it...?-----v	311
CONSENSUS-F	-----e.dii---q--sd-----Nes-----?-----I?-----r...--	298

HIV1 ENV CONSENSUS

CONSENSUS-G	----?-----e.-I-----?d---v----n-sie?------i?f.....	297
CONSENSUS-H	. -?-----?---D-T-N----K-----?----.----.I?-----?---	38
CONSENSUS-O	--I---T-Skg.kIr-Mgk--?dsg-N---T-N?-i-mt-e--.g?-v.Qe...i?----.m..-W	277
CONSENSUS-U	-----E.-?i-----d---?----ne?-----?-----?-----?-----?	269
CONSENSUS-CPZ	?--?---????--?---????K?????V?????E??-??-.G-?-?.?....Q-----M..T-	140
V3 neutralization loop ->		
	* ^ ^ ^ ^ ^ ^ ^ ^ ^	
CONSENSUS-A	ya.tgdi.....iG.dirqAhCnvsr?eWn?tlq?v....a?qLr?..yf??nkt....?iif?n.ss	319
CONSENSUS-B	-t.--?---?...---i---ak---n---kqi....v-k---e?q?-----?----v-nq?--	340
CONSENSUS-C	---.----i---i?k----?kk-ae..h-p-----?----?	325
CONSENSUS-D	-t?---pr????..?----i-a?---k---q----k-gd?.l1.....t---kp.--	327
CONSENSUS-E	-r.-----k-y-EINGTk---e?-kq----tek-ke..H--.n-----?qP?p-	358
CONSENSUS-F	---?----k----gtq----e----?ak-ks..h----?----k-ns.--	340
CONSENSUS-G	---?----?----?----?em---n----?----?----?----?----t-ns?a	332
CONSENSUS-H	?-.?----?----?----?----I?----?----?----?----?H?----?----?----?P.--	62
CONSENSUS-O	-S.M..1???n?k???s?-Y?YnaTd?ka-kqt....eRYLe..Lv....?----?vtm?-n?s?-	318
CONSENSUS-U	---?----i---t?----q----?----?----?----?----?----?----?	304
CONSENSUS-CPZ	-N.?E??....?---T?---?---N?T?---?----?----?----?----?A-?---?..?	155
CD4		
	CD4 * * ^ ^ ^ ^ ^ ^ ^ ^	
CONSENSUS-A	GGD1EitthsfnCggef?FYCnts?lF.nstw?????????....n?t?????????n?t????..?n??n	353
CONSENSUS-B	---p---vm-----tq-----?----?----?----?----?----?----?	374
CONSENSUS-C	-----r-----?----y-----?----?----?----?----?----?	357
CONSENSUS-D	---p-----?----?----?----?----?----?----?----?	358
CONSENSUS-E	-----m-h---r----t---n-cig-----e-m----gc..-g-----?	398
CONSENSUS-F	-----m---r----?----d?----?----?----?----?----?	370
CONSENSUS-G	-----r----g---?----?s?----?----?----?----?----?	363
CONSENSUS-H	---?----?----?----?K----?----?----N----.G-----	91
CONSENSUS-O	---?----v-hlh---H----?----m..-y-Fsc----?----?----n-----	354
CONSENSUS-U	-----t-----?----?----?----?----?----?----?	332
CONSENSUS-CPZ	?--P-V??-?----?----?----?----?----?I-----	174
	*	CD4 * ^ ^ ^ CD4 ^ ^ ^
CONSENSUS-A	??It..lq.CrI.kqivnm.wQrvqq.amYapPI?.g?irc?snITGllLTrDgg??....nns???.?	399
CONSENSUS-B	---?----p----?----i----e---k----?----q---s-----?----?----?----?----?	420
CONSENSUS-C	---.----p----?----i----e---r?----?----n-t-k-----?----?----?t----?----?	404
CONSENSUS-D	-t----p----?----i----?----g---k----e----?----s-----?----?----?----?	405
CONSENSUS-E	---.----P----k----?----i----ga----?----s..r-n-V-----I-----a?----?----t----?	448
CONSENSUS-F	-t----p----?----e---r----a---a---n-t-n-----?----?----?----t----?----?	419
CONSENSUS-G	---.----p----?----r----?----A?----n-t-----?----?----t----?	410
CONSENSUS-H	---?----?----?----?----?----K-T-----I--S-HV.....?----AE....	134
CONSENSUS-O	gT??.iP.--L.r-v-Rs.-m-G-S.gl----?----NLt-----MI-Q1?..pW....s----	401
CONSENSUS-U	---?----P----?----r----?----?----?----?----n?----?----?----?	375
CONSENSUS-CPZ	G?I..?----?----?----?----?R.-N?----?----?----?----?----?----?	198
	^ ^ ^ gp120 / gp41	
CONSENSUS-A	?---?netFrPgGgdmrdNWrsSELYkYKvVkiePlGvaPtr.akrRvV....eREKRA?vg.lGavflgf	457
CONSENSUS-B	---t-i-----?----k-----?----q-----?i----m-----	478
CONSENSUS-C	---?----?----e-k-----?----?----?----?----?----?----?	460
CONSENSUS-D	-----r-----?----?----?----?----?----?----?----?	463
CONSENSUS-E	-----NiK-----Q-----i-----?----?----?----?----?----?	506
CONSENSUS-F	...?----n-k-----e-----?----q-----k-----?----l-----	475
CONSENSUS-G	-----k-----?----R-----g-----?----?----?----?	468
CONSENSUS-H	...?V-----?----?----?----?----?----?----?----?----?	184
CONSENSUS-O	.----?----l----?----k-I---T---f----rvK-FS----ki-RP?Igt?t?H-----?----ML---v	459
CONSENSUS-U	...?----?----?----?----?----?----?----?----?----?----?----?	431
CONSENSUS-CPZ	...?----?----?----?----?----?----?----?----?----?----?	225
CONSENSUS-A	lGaAGstmgaasitltvqarql1sgivqqqnllraiaeaqgh11kltVGWIKQLQARVLAvErYLrD.QQ	526
CONSENSUS-B	-----?----?----n-----?----q-----?----k?--	546
CONSENSUS-C	-----?----?----m-Q-----t----i----k----?	529
CONSENSUS-D	-----?----?----N-----?----k-----?----K----?	531
CONSENSUS-E	-----?----?----Q-----?----K----.K	575
CONSENSUS-F	-----?----n-----?----Q-----?----k----?	544
CONSENSUS-G	-----?----V-----?----Q-----?----?----?----?	536
CONSENSUS-H	-----?----?----?----?----?	227
CONSENSUS-O	-S----?Ata---tht-?K----D----Q----?----R-S----R---R-L--L-TliQN.--	526
CONSENSUS-U	-----?----?----?----?----Q-----?----?----?----?	495
CONSENSUS-CPZ	?----?----?----?----?----?----?----Q-S?--V----?----?----?----?	278

	* * ^^	^^	^^	^^	
CONSENSUS-A	LLGIWCGSGK1CtTnVPWNsSW.....S.Nks??dIWdnMTWlqWdKEisnYT?iIY?.LiEesq				581
CONSENSUS-B	-----a-a-----.-?---l---?---me-er--d---l-t-----				601
CONSENSUS-C	-----a-----.-?---q-----m---r-----t--r?-l-d--				586
CONSENSUS-D	-----h-----.-.r-L-e--?---mE-ER--d---G1-s-----				588
CONSENSUS-E	f--L----I----A----t-----.-.r-fEE-n---iE-eR-----nq--e.ILT--				634
CONSENSUS-F	---L-----t-----.-.qEe-g---ME-e---SnE--R---?--				602
CONSENSUS-G	-----t-----.-.fnE-----Ie-eR--N---q--n.-l----				594
CONSENSUS-O	--nL---K---?---Y-S-k---kt-?g.....?---?dnes---L-Q?--qq-n?SSt---.e-QkA-				579
CONSENSUS-U	--?---?---?---?-----.-?---?---?---?ME-?R-?---V---.-?---				536
CONSENSUS-CPZ	?-?L---?---?---?T---N---??????????.?---?---?Q---?LV?---?G---?---?L??A?				310
			\ / 3'sj		
CONSENSUS-A	nqQEkdLLaLDkWanLwnWFdIsnWLWYIriFimIVGGLIGLRivfaVlsiInRVRqGYSP1SFQt1				651
CONSENSUS-B	-----e-e----s-----?t----k-----v-----v-----?-----				669
CONSENSUS-C	-----k-----q-----t?----k-----i-----v-----				655
CONSENSUS-D	?-----e---?---S---s-T?----k-----lV-----				655
CONSENSUS-E	---DR--K---e---S-----T-----K-----i-----V-----p				704
CONSENSUS-F	-----e-----S-----K-----V-----K-----l-h				672
CONSENSUS-G	?-----?S---s-s-----k-----V-----				662
CONSENSUS-O	v---?---kk---E---E---Si---l---TK---K-A-I---A---?---?V?MI?---N1VkNI---Q---L-iP				644
CONSENSUS-U	?---?---?---?-----S-----?---?-----K-----?---?---?-----				593
CONSENSUS-CPZ	?---?---?---?E---?---S-----T-----K---?---?I?-----?-----?R?-----?---?---				355
	-< tat cds	*			
CONSENSUS-A	tp?pr?pdRperIeeeGG.eQdrdRSirLVsGFLalAWDDLRLsLCLfsYHrLRDfill.a.ar.tVell				714
CONSENSUS-B	1-a--g-----g-----.-r-----g---?-----i-----Ll---?v?---?i---				733
CONSENSUS-C	--n---g---lg?-----.				719
CONSENSUS-D	1-a--g-----g-----.-G-----n---s---i---n-----l---?---.I---				720
CONSENSUS-E	?HhQ-E-----g---g-----V-----				768
CONSENSUS-F	i-S---E-----g---G-----Gk---V---n-----v-----N-----rH-----?---.I-?..				734
CONSENSUS-G	?HHQ-e-----g---G-----				726
CONSENSUS-O	i?hq?E.a?T-G-TG-g---e---p-w?p?Pq---p-LYT---TII-W?--L-SNLaSg.I.....				696
CONSENSUS-U	?---G---G?-----.-?---?---?-----?-----?-----?-----L?--.?.??.??.?-----				636
CONSENSUS-CPZ	?---Q?-.?----?E---?---.?---?---?---?---?-----N-GIW---QS-TSLACN.V.W-.QLKT-				396
	-< rev cds	^^			
CONSENSUS-A	ghsslkglrlg.....weglkYL.wNLllyWgrELK?SAinLldtiAiavAgwtDrvIEigQrigRAi				775
CONSENSUS-B	-rr?....?---a---w---q---sq---n---vs---nat---Eg-----vv---a?---				787
CONSENSUS-C	-r---r---qr-----?-----gs-vq---l---k---s---?-----Eg---i---?i---?---				777
CONSENSUS-D	-R.....R-----a-----q---?q---n---S-----Eg---?---?v---a?--v				771
CONSENSUS-E	-----R-----G-----Q---I---S---nat-----vA-gaW---				830
CONSENSUS-F?R---R-----A---l---G---t---Q---N---s---N---T---v---Eg-----AL---a---				791
CONSENSUS-G	-Rni-----?-----q---N-----?-----N-----vv---aC---				787
CONSENSUS-O	.qklIs?-g--LWI1GQk?IeaCR?caAvtQ--LQ--qn-T?---l-V---N---gi-lGi-----G-				761
CONSENSUS-U	?-----?-----?-----?-----?-----G?---V?---?-----S?AT---?---EG---?-----?-----?---				675
CONSENSUS-CPZ	--LI-HS---L.....R-R-CL-.GGIIQ---K---I---S---AT---EG---I--AF-VTL-I-				458
CONSENSUS-A	lnIPrRIRQGlEraL\$				791
CONSENSUS-B	-h---?-----				802
CONSENSUS-C	?-----f-a---q-				792
CONSENSUS-D	-h---?-----				786
CONSENSUS-E	-h-----				851
CONSENSUS-F	-----				807
CONSENSUS-G	-----				803
CONSENSUS-O	?-----S---				776
CONSENSUS-U	?-----F---?---?-				688
CONSENSUS-CPZ	R-----				474